Grundlagen der asymptotischen Verteilungstheorie

Die asymptotische Verteilungstheorie beschäftigt sich mit der Verteilung statistischer Kenngrößen, wie z.B. dem empirischen Mittel, wenn die Anzahl der Beobachtungen immer größer wird (gegen unendlich geht). Die asymptotischen Verteilungen sind oft leichter zu bestimmen als die für endliche Stichproben. Gleichzeitig geben sie oft eine sehr gute Approximation auch für Stichproben mit endlich vielen Beobachtungen.

Definition: Stochastische Konvergenz

$S_1, S_2, \ldots, S_n, \ldots$ sei eine Folge von Zufallsvariablen. Beilspielsweise kann S_n der empirische Mittelwert aus einer Stichprobe mit n Beobachtungen sein. Die Sequenz der Zufallsvariablen $\{S_n\}$ konvergiert stochastisch gegen den Grenzwert μ, wenn gilt

$$\lim_{n\to\infty} Pr\left(|S_n - \mu| \geq \delta\right) = 0$$

für beliebiges $\delta > 0$. Man schreibt dann oft $S_n \xrightarrow{p} \mu$ oder $\operatorname{plim} S_n = \mu$. S_n konvergiert also stochastisch gegen μ, wenn die Wahrscheinlichkeit, dass die absolute Differenz $|S_n - \mu|$ größer als ein beliebiges δ, für n gegen unedlich, gegen null geht.

Gesetz der Großen Zahlen

Das Gesetz der großen Zahlen besagt, dass unter bestimmten Bedingungen der empirische Mittelwert aus einer Stichprobe gegen den Erwartungswert konvergiert.
Wenn $y_1, \ldots y_n$ unabhängig und identisch verteilte Zufallsvariablen mit $E(y_i) = \mu$ und $Var(y_i) = \sigma^2 < \infty$ sind, dann gilt

$$\operatorname{plim} \bar{y} = \operatorname{plim}\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) = \mu$$

Definition: Konvergenz in Verteilung

$S_1, S_2, \ldots, S_n, \ldots$ sei eine Folge von Zufallsvariablen mit den jeweiligen Verteilungsfunktionen $F_1, F_2, \ldots, F_n, \ldots$. Ferner sei S eine Zufallsvariable mit Verteilungsfunktion F. Die Folge der Zufallsvariablen $\{S_n\}$ konvergiert in Verteilung gegen die Zufallsvariable S, wenn gilt

$$\lim_{n\to\infty} F_n(x) = F(x)$$

an allen Stellen x, an denen $F(x)$ stetig ist. Man schreibt oft $F_n \xrightarrow{d} F$.

Zentraler Grenzwertsatz

Zentrale Grenzwertsätze geben Bedingungen an, unter denen Funktionen von empirischen Mittelwerten in Verteilung gegen eine normalverteilte Zufallsvariable konvergieren. Die Einfachste Version besagt:

Wenn $y_1, \ldots y_n$ unabhängig und identisch mit $E(y_i) = \mu$ und $Var(y_i) = \sigma^2 < \infty$ verteilt sind, gilt

$$\frac{\bar{y} - \mu}{\sigma_{\bar{y}}} \xrightarrow{d} N(0,1)$$

Dabei ist $\bar{y}$ der empirische Mittelwert und $\sigma_{\bar{y}}$ die Standardabweichung von $\bar{y}$. Da $\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}}$, kann der Ausdruck umgeschrieben werden zu

$$\begin{aligned} \frac{\sqrt{n}(\bar{y} - \mu)}{\sigma} &\xrightarrow{d} N(0,1) \\ \text{oder:} \quad \sqrt{n}(\bar{y} - \mu) &\xrightarrow{d} N\left(0, \sigma^2\right) \\ \text{oder:} \quad \frac{\sum_{i=1}^{n}(y_i - \mu)}{\sqrt{n}} &\xrightarrow{d} N\left(0, \sigma^2\right) \end{aligned}$$

Continous Mapping Theorem

$$S_n \xrightarrow{p} \mu \quad \Rightarrow \quad a(S_n) \xrightarrow{p} a(\mu)$$

Dabei ist $a(\cdot)$ eine stetige Funktion. Das impliziert für $S_n \xrightarrow{p} \mu_1$ und $x_n \xrightarrow{p} \mu_2$:

$$S_n + X_n \xrightarrow{p} \mu_1 + \mu_2$$

$$S_n \cdot X_n \xrightarrow{p} \mu_1 \cdot \mu_2$$

$$\frac{S_n}{X_n} \xrightarrow{p} \frac{\mu_1}{\mu_2}, \text{ wenn } \mu_2 \neq 0$$

Slutzky Theorem

$$x_n \xrightarrow{p} \mu \text{ und } S_n \xrightarrow{d} S \quad \Rightarrow \quad x_n S_n \xrightarrow{d} \mu S$$

Beachte: μ ist eine Konstante und S eine Zufallsvariable.

Erweiterung auf Vektoren und Matrizen

Alle Definitionen, Sätze und Theoreme haben wir bisher für skalare Zufallsvariablen angegeben. Sie lassen sich aber leicht auf Vektoren und Matrizen von Zufallsvariablen erweitern. Das soll für das Slutzky Theorem genauer erläutert werden. Es gelte

$$\boldsymbol{A}_n \xrightarrow{p} \boldsymbol{A} \text{ und } \boldsymbol{s}_n \xrightarrow{d} \boldsymbol{s}$$

$\boldsymbol{A}_n$ und $\boldsymbol{A}$ seinen $(k \times k)$-Matrizen wobei $\boldsymbol{A}$ eine feste Matrix ist. $\boldsymbol{s}_n$ und $\boldsymbol{s}$ seien $(k \times 1)$-Vektoren. Dann gilt:

$$\boldsymbol{A}_n \boldsymbol{s}_n \xrightarrow{d} \boldsymbol{A}\boldsymbol{s}$$

Für Erwartungswert und Varianz von $\boldsymbol{A}\boldsymbol{s}$ gilt

$$E(\boldsymbol{A}\boldsymbol{s}) = \boldsymbol{A}E(\boldsymbol{s})$$

$$
\begin{aligned}
Var\left(\boldsymbol{A}\boldsymbol{s}\right) &= E\left[\left(\boldsymbol{A}\boldsymbol{s}\right)\left(\boldsymbol{A}\boldsymbol{s}\right)'\right] - E\left(\boldsymbol{A}\boldsymbol{s}\right)E\left(\boldsymbol{A}\boldsymbol{s}\right)' \\
&= \boldsymbol{A}E\left(\boldsymbol{s}\boldsymbol{s}'\right)\boldsymbol{A}' - \boldsymbol{A}E\left(\boldsymbol{s}\right)E\left(\boldsymbol{s}'\right)\boldsymbol{A}'
\end{aligned}
$$

Dabei bezeichnet $Var\left(\boldsymbol{A}\boldsymbol{s}\right)$ die Varianz-Kovarianz-Matrix des Vektors $\boldsymbol{A}\boldsymbol{s}$. Ist $E\left(\boldsymbol{s}\right) = 0$ gilt

$$
\begin{aligned}
Var\left(\boldsymbol{A}\boldsymbol{s}\right) &= \boldsymbol{A}E\left(\boldsymbol{s}\boldsymbol{s}'\right)\boldsymbol{A}' \\
&= \boldsymbol{A}Var\left(\boldsymbol{s}\right)\boldsymbol{A}'
\end{aligned}
$$

mit $Var\left(\boldsymbol{s}\right)$ der Varianz-Kovarianz-Matrix von $\boldsymbol{s}$.